Approximation of genetic code via cell-free protein synthesis directed by template RNA

MARSHALL W. NIRENBERG, J. HEINRICH MATTHAEI,[1] OLIVER W. JONES, ROBERT G. MARTIN, AND SAMUEL H. BARONDES

National Heart Institute, National Institutes of Health, Bethesda, Maryland

ALTHOUGH THE FIELD OF GENETICS attained a high degree of sophistication years ago, the molecular basis of genetic information storage and retrieval lagged behind. Only within the last 10 years has the chemical structure of DNA been established, mainly through the work of Chargaff et al. (3), Wilkins et al. (29), and Watson and Crick (28). Some 8 years ago Gamow (5) proposed a theoretical code and thereby stimulated a great deal of interest in this problem. The direct biochemical approach, that is, comparison of the nucleotide sequence of a gene with the amino acid sequence of its corresponding protein, posed technical problems of great magnitude. Some of these difficulties have been circumvented by the experimental approach summarized here.

DEVELOPMENT OF CELL-FREE ASSAY FOR MESSENGER RNA

C^{14}-amino acid incorporation into protein was studied in a cell-free *E. coli* system patterned generally after the systems established by Tissières, Schlessinger, and Gros (25), also by Lamborg and Zamecnik (9) and by Spiegelman and his associates (24). A major difficulty in the study of cell-free protein synthesis in *E. coli* systems had been the necessity for preparing fresh enzyme extracts for each experiment. Techniques were not available for stabilization and storage of these enzyme extracts comparable to those available for mammalian systems. *E. coli* extracts were found to be stable in the presence of mercaptoethanol and could be dialyzed and stored frozen without undue loss of activity (15).

Independently, Tissières et al. (25), Kameyama and Novelli (8), and Nisman and Fukuhara (20) reported an inhibition of amino acid incorporation by deoxyribonuclease. We also observed such inhibition and studied its characteristics (15), for here appeared to be a cell-free system, in which DNA directed the synthesis of protein. The effect of deoxyribonuclease on C^{14}-valine incorporation is shown in Fig. 1. In the absence of deoxyribonuclease, C^{14}-valine was rapidly incorporated into protein. Deoxyribonuclease completely inhibited incorporation occurring after 30 min of incubation (15). This type of experiment suggested that the deoxyribonuclease-inhibited system might be dependent on the availability of messenger RNA. Crude messenger RNA fractions were found to stimulate amino acid incorporation and proved to be a requirement for cell-free protein synthesis (17, 18). This technique afforded a highly sensitive, cell-free assay for messenger RNA and provided the rationale for all of our subsequent work.

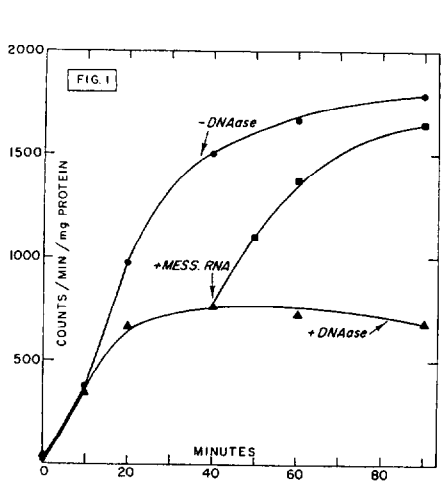
[1] NATO Postdoctoral Fellow.

EFFECT OF SYNTHETIC POLYNUCLEOTIDES IN DIRECTING PROTEIN SYNTHESIS

Discovery of polynucleotide phosphorylase by Grunberg-Manago and Ochoa (7) made available to us chemically defined polyribonucleotides which were assayed for messenger activity in the cell-free *E. coli* system (17, 18). Figure 2 shows the effect of the synthetic polynucleotide, polyuridylic acid,[2] on incorporation of C^{14}-phenylalanine. In the absence of poly U, very little phenylalanine was incorporated into protein. Addition of 10 μg of poly U to a 1-ml reaction mixture resulted in a 1,000-fold increase in phenylalanine incorporation. The synthesized protein was purified by extraction of precipitated washed protein with 33% HBr in glacial acetic acid followed by precipitation with H_2O. The synthesized protein was shown to have unusual solubility and stability properties characteristic of authentic polyphenylalanine (18). The effect of poly U in directing the incorporation of 18 other C^{14}-amino acids into protein was studied. Marked specificity for phenylalanine was observed; however, a small stimulation of leucine incorporation was observed (3% that of phenylalanine) (14).

The data of Fig. 3 show that the template activity of poly U was dependent on its molecular weight (14).

[2] The following abbreviations are used: poly U, polyuridylic acid; poly A, polyadenylic acid; octa A, octaadenylic acid; tetra A, tetraadenylic acid; poly C, polycytidylic acid; TMV, tobacco mosaic virus.



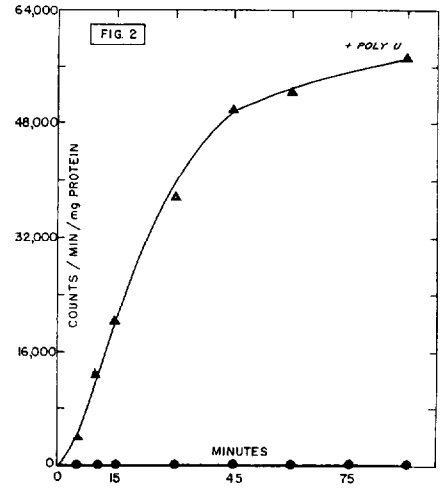
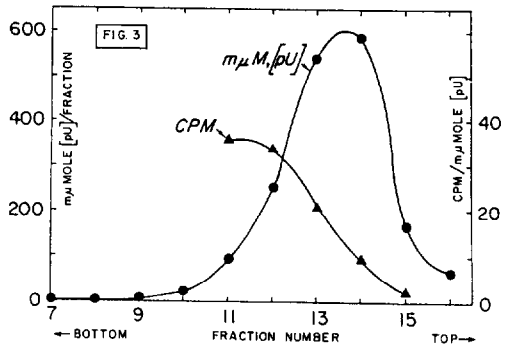


FIG. 1. Cell-free *E. coli* assay for messenger RNA: ● Minus deoxyribonuclease, ▲ 10 μg deoxyribonuclease, ■ 500 μg of crude messenger RNA prepared from yeast ribosomes added at indicated time. For details see refs. 15 and 18.

FIG. 2. Stimulation of C^{14}-L-phenylalanine incorporation by polyuridylic acid. ● Minus polyuridylic acid, ▲ 10 μg polyuridylic acid added. For details see ref. 18.



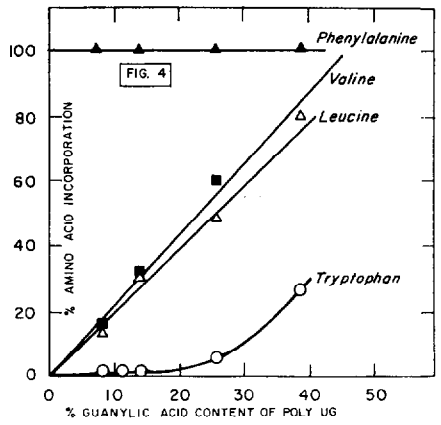


FIG. 3. Relationship between molecular weight of polyuridylic acid and its activity in stimulating C^{14}-phenylalanine incorporation into protein. Polyuridylic acid was separated into fractions of different sizes by sucrose density gradient centrifugation. Each fraction was then assayed for its ability to stimulate C^{14}-phenylalanine incorporation. For details see ref. 14.

FIG. 4. Relationship between guanylic acid content of different poly UG preparations and their their abilities to direct C^{14}-valine, C^{14}-leucine, and C^{14}-tryptophan incorporation into protein relative to C^{14}-phenylalanine incorporation.

Larger molecules of poly U were separated from smaller molecules by centrifugation through a sucrose gradient. The heavier poly U molecules were considerably more active than the lighter fractions. Poly U of molecular weight of 50,000–100,000 was very active in this system; however, the minimum molecular weight necessary for activity is unknown.

Stoichiometry experiments demonstrated that phenylalanine incorporation was proportional to the concentration of poly U in the proper range. The data of Table 1 demonstrate the quantitative relationship between phenylalanine incorporated and poly U present in reaction mixtures. Results of three different experiments are presented. About 1.5 mμmoles of uridylic acid residue in poly U were required to direct incorporation of 1.0 mμmole of phenylalanine (14). These data alone cannot be used to determine the number of uridylic acid residues in a phenylalanine coding unit, for additional results, discussed later, suggest that a molecule of poly U may function catalytically, by directing the synthesis of a number of molecules of polyphenylalanine.

Since one or more uridylic acid residues in poly U appeared to be the RNA code word which corresponded to phenylalanine, application of this experimental technique to determine other code words and the general

TABLE 1. *Relationship between C^{14}-L-phenylalanine incorporated and polyuridylic acid present*

Exp. No.	mμmole pU (in Polyuridylic Acid)	mμmole C^{14}-L-phenylalanine	$\frac{\text{mμmole pU}}{\text{mμmole } C^{14}\text{-phenylalanine}}$
1	11.35	9.12	1.24
2	17.7	10.25	1.73
3	25.7	16.9	1.52

In each experiment the amount of polyuridylic acid present was limiting. Final volume was 0.5 ml. Reaction mixtures were incubated for 60 min at 37 C as described elsewhere (14).

characteristics of the genetic code appeared reasonable (17, 18). Randomly ordered polynucleotides containing different bases were tested by Ochoa and his collaborators (12, 10, 11, 22, 23) and by ourselves (13, 14) to see if such polymers would code for different amino acids. In Table 2 are presented the effects of randomly ordered polynucleotides on amino acid incorporations. The composition of the different polynucleotides was obtained by base-ratio analysis. Since each preparation of polynucleotide differs in molecular weight, it is difficult to compare directly the activity of one polynucleotide with another. Therefore, the figures in the main part of the table represent the per cent of an amino acid incor-

TABLE 2. *Amino acid incorporation into protein stimulated by randomly mixed polynucleotides*

Amino Acid	Polynucleotides and Base Ratios					
	UA U = 0.87 A = 0.13	UC U = 0.39 C = 0.61	UG U = 0.76 G = 0.24	UAC U = 0.834 A = 0.050 C = 0.116	UGC U = 0.341 G = 0.152 C = 0.502	UGA U = 0.675 G = 0.291 A = 0.034
Phenylalanine	100	100	100	100	100	100
Arginine	0	0	1.1	0	49.3	2.9
Alanine	1.9	0	0	1.0	40.4	0.9
Serine	0.4	160	3.2	3.6	170	2.3
Proline	0	285	0	0	188	0
Tyrosine	13	0	0	8	1.0	8.6
Isoleucine	12	1.0	1.0	4.8	5.4	8.4
Valine	0.6	0	37	0.4	29.8	75
Leucine	4.9	79	36	5.1	157	44
Cysteine	4.9	0	35	0	5.4	46
Tryptophan	1.1	0	14	0	1.6	23
Glycine	4.7	0	12	0.5	9.7	15
Methionine	0.6	0	0	0.6	1.5	8
Glutamic acid	1.5	0	0	1.2	0.44	6.2

Figures in main part of the table represent incorporation of any amino acid compared to phenylalanine incorporation expressed as percentages (mμmoles amino acid incorporated/mμmoles phenylalanine incorporated × 100). Approximately 25 μg of each polynucleotide was added to each 0.5-ml reaction mixture; these are described elsewhere (14). Samples were incubated at 37 C for 15 min. Incorporation of C^{14}-phenylalanine in counts per minute due to the addition of polynucleotides UA, UC, UG, UAC, UGC, and UGA were 731, 2,900, 714, 804, 2,144, and 2,744, respectively. The reproducibility of the above percentage figures was ±3.

porated compared to phenylalanine incorporation stimulated by the same polynucleotide. Marked polynucleotide specificity in stimulating incorporation is apparent.

A degenerate code is one in which two or more different coding units can direct the incorporation of the same amino acid into protein. Since leucine can be coded by both poly UC and poly UG, the code, with respect to leucine, is degenerate (13, 14).

Different preparations of poly UG containing increasing percentages of G were synthesized and assayed (Fig. 4). The relative amounts of valine, leucine, and tryptophan incorporations, compared to phenylalanine incorporation, was determined for each polynucleotide. Leucine and valine incorporation was proportional to the amount of G in poly UG. However, significant tryptophan incorporation occurred only when a high proportion of G was present. These data suggest that the coding units for both leucine and valine contain one G, that the coding unit for tryptophan contains two G's, and that each preparation of poly UG contained equal numbers of coding units for leucine and valine. Similar experiments were performed with other amino acids and quantitative estimates were obtained of the number of G, C, or A residues per coding unit.

These results are summarized in Table 3. The probability of any sequence of three nucleotides occurring in randomly ordered polynucleotide of analyzed base ratio can be calculated and may be compared with the probability of obtaining a sequence of UUU. The theoretical probabilities of obtaining such triplets, relative to UUU, are compared in Table 3 with the experimentally obtained data. Data obtained experimentally agree strikingly well with theoretical predictions. Such quantitative data demonstrate that the number of G's, A's, or C's in coding units can be determined experimentally.

The sequence of nucleotides in each coding unit is not known. The current status of the code is similar to that of an anagram, where the letters of coding units have been determined, but the sequence is unknown.

Although the uridylic acid content of RNA viruses is not excessive, a surprisingly high proportion of U has been found in coding units thus far. This dichotomy cannot be explained at the present time. However, it is probable that only coding units containing U have been selected for by the assay method and that additional degeneracies without U will be found.

The question may be asked, "Do the code words agree with known genetic data?" In Table 4 coding results are compared with nitrous acid-induced amino acid substitutions in tobacco mosaic virus protein. Only amino acid substitutions found more than once have been cited. TMV protein contains 158 amino acids, and the entire sequence is known. Treatment of RNA with nitrous acid results in deamination of nucleotides (12) and hence mutations (6). Most nitrous acid-induced amino acid replacements should be due to the replace-

TABLE 3. *Summary of code words**

Amino Acid	Coding Unit Composition	Theoretical	Experimental
Phenylalanine	UUU...	100	100
Arginine	UCG...	67	49
Alanine	UCG...	67	40
Serine	UUC...	157	160
	(UCG)		
Proline	UCC...	244	285
Tyrosine	UUA...	13	13
Isoleucine	UUA...	13	12
Valine	UUG...	32	37
Leucine	UUG...	32	36
	(UUC)	157	79
Cysteine	UUG...	10.6	10.6
	(UGG)?		
Tryptophan	UGG...	11	14
Glycine	UGG...	11	12
Methionine	UGA...	2	8
Glutamic acid	UGA...	2	6
Lysine	UAA...	?	?

* Sequence of nucleotides in code words is not specified. Coding ratio is not known definitively. However, assuming a triplet code, the probability of a triplet occurring in a polynucleotide, relative to UUU, may be calculated from the base-ratio data presented in Table 2. For example, if poly UG had a base ratio of 3U/1G, the probability of obtaining the sequence UUU would be $\frac{3}{4} \times \frac{3}{4} \times \frac{3}{4} = \frac{27}{64}$. The probability of obtaining the sequence UUG would be $\frac{3}{4} \times \frac{3}{4} \times \frac{1}{4} = \frac{9}{64}$. Thus, 3 UUU would occur for 1 UUG and, assuming the frequency of UUU to be 100%, the frequency of UUG would be 33%. Dots after each "word" are used to indicate the possible presence of additional uridylic acid residues.

TABLE 4. *Comparison of nitrous acid-induced replacements in tobacco mosaic virus protein with the nucleotide composition of RNA coding units*

No. of Mutant Strains*		Amino Acid Replacement	Nucleotide Composition of Corresponding Coding Units	Possible Nucleotide Changes
A	B			
1	2	Ser	UUC...	C
		↓	↓	↓
		Phe	UUU...	U
1	2	Glu	UAG...	A
		↓	↓	↓
		Gly	UGG...	G
1	2	Prol	UCC...	C
		↓	↓	↓
		Leu	UUC...	U
	2	Isoleu	UUA...	A
		↓	↓	↓
		Val	UUG...	G
5		Arg	UCG...	C
		↓	↓	↓
		Gly	UGG...	G

* The mutant amino acid replacement data cited and obtained either by A, Tsugita and Fraenkel-Conrat (30), or B, Wittmann (26).

ment, during virus replication, of either a cytidine by a uridine or an adenine by a guanine (4).

All of the amino acid replacements with the exception of the last are the result of a conversion of either a C to a U or an A to a G. Such results demonstrate confirmation between amino acid replacement data and proposed nucleotide compositions of coding units (14). Amino acid replacements in corticotropins, insulins, and cytochromes of different species have been examined also. Over 90% of these acid substitutions corresponded to a replacement of only one nucleotide per coding unit. Such data suggest that at least part of the code may be universal.

Nonoverlapping and overlapping codes would be read as follows:

UUUUUU	UUUUUUU
Nonoverlapping	Overlapping

Brenner on theoretical grounds ruled out overlapping triplet codes because of the restrictions such codes would impose on nearest-neighbor frequencies of amino acids (2). If the code were overlapping and one nucleotide in a coding unit were replaced with another due to mutation, amino acid substitutions might be expected to occur in clusters. The bulk of amino acid substitutions found in TMV protein do not occur in clusters, and these data suggest that the code is of the nonoverlapping type (26, 30, and personal communication from Tsugita and Fraenkel-Conrat).

A number of polynucleotides which do not contain U have been tested (Table 5). Poly AC at the indicated base ratios determined by analysis, were found to direct small amounts of threonine and proline into protein. Non-U-containing polynucleotides of different base ratios, or more sensitive assays, may be necessary to demonstrate additional degeneracies.

When poly U is mixed with poly A, double- and triple-stranded helices are formed which are completely inactive in directing polyphenylalanine synthesis. The effects of oligo A and poly A on polyphenylalanine synthesis are presented in Fig. 5. The stability of the oligo A-poly U helix is influenced strongly by the chain length of oligo A. Thus octa A forms more stable helices with poly U than tetra A. It may be seen that octa A and poly A are very effective inhibitors of polyphenylalanine synthesis. In the reaction mixtures used, triple-stranded helices would be formed (U-A-U). A 99% inhibition of polyphenylalanine synthesis was observed at stoichiometric poly A concentrations, that is, 2 U's in poly U per 1 A in poly A. In contrast, addition of poly C had little effect on polyphenylalanine synthesis. Many mechanisms for enzyme repression can be envisioned with strands of RNA base pairing with parts of DNA or complimentary RNA. It will be of critical importance to determine whether the type of repression demonstrated in vitro with poly U and poly A may occur also in vivo.

FATE OF MESSENGER RNA

H^3-poly U was prepared and its fate in reaction mixtures was determined (1). Sucrose density gradient centrifugation experiments revealed that H^3-poly U became associated with a polydisperse peak of ribosomes which sedimented faster than 70S ribosomes. Almost no poly U was found on 70S ribosomes. After 3 min of incubation approximately 90% of the poly U disappeared and was recovered as mononucleotides. The rest of the poly U still was associated with 100S ribosomes. After 10 min of incubation most of the poly U had been hydrolyzed to mononucleotides (90% were 5′-mononucleotides, 10% were 2′, 3′-cyclic mononucleotides or 3′-mononucleotides.) These data show that poly U is bound to 100S rather than 70S ribosomes. In addition, the observed rapid breakdown of poly U suggests that

TABLE 5. *Effects of non-U-containing polynucleotides on amino acid incorporation*

C^{14}-L-Amino Acid	Minus Poly-nucleotide	AG 7:1	AG 1.5:1	AC 5:1	AC .05:1
Phenylalanine	10	10	9	10	10
Valine	10	9	8	10	8
Serine	20	20	20	15	18
Methionine	10	10	9	10	11
Histidine	10	5	11	10	10
Leucine	30	30	25	28	30
Lysine	10	10	15	13	10
Aspartic acid	10	10	10	12	10
Alanine	10	10	11	10	10
Cysteine	20	10	10	20	10
Glycine	10	10	11	10	10
Glutamic acid	10	20	10	10	20
Proline	10	10	11	10	70
Isoleucine	10	10	10	10	11
Tyrosine	10	10	8	10	9
Threonine	10	10	9	30	10
Arginine	10	10	9	8	8

Figures represent the incorporation of C^{14}-amino acids in $\mu\mu$moles. Base ratios were obtained by analysis.

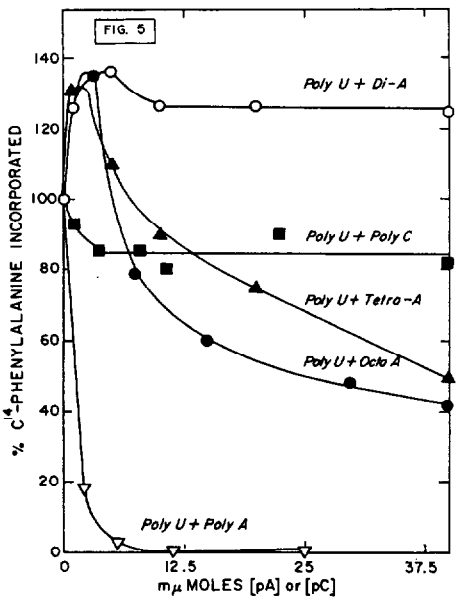


FIG. 5. Repression of polyphenylalanine synthesis by oligo and poly A. 12.5 mμmoles of uridylic acid residues in poly U and the indicated amount of oligo A, poly A, or poly C were incubated at 24C for 15 min before adding to reaction mixtures. 12.5 mμmoles of uridylic acid in poly U alone stimulated incorporation of 2,250 counts/min of C^{14}-phenylalanine into protein. Thus 2,250 counts/min represents 100% incorporation.

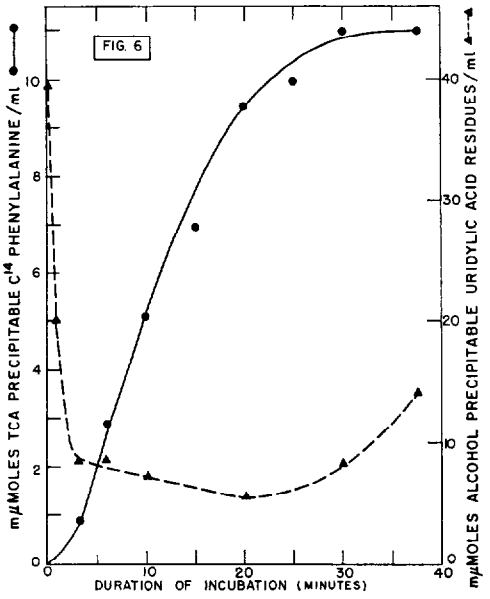


FIG. 6. Comparison between rate of polyphenylalanine synthesis and rate of H^3-poly U degradation. Reaction mixtures used for assay of precipitable H^3-poly U contained H^3-poly U and C^{12}-phenylalanine. Reaction mixtures used for assay of phenylalanine incorporation contained C^{14}-phenylalanine. The 2 reaction mixtures, each in a total volume of 2 ml, were incubated for 40 min and 0.2-ml aliquots were removed at the indicated time intervals and were transferred to either 2 ml of 10% trichloracetic acid at 4C for assay of phenylalanine incorporation into protein or into 0.2 ml of magnesium acetate and 0.4 ml of absolute ethanol at 4C for analysis of precipitable H^3-poly U (1).

only the remaining 10–20% directs protein synthesis. These data and the stoichiometry data of Table 1 suggest that poly U functions catalytically; that is, each poly U molecule directs the synthesis of more than one polyphenylalanine molecule (14, 1).

Figure 8 illustrates an experiment performed with unlabeled poly U and C^{14}-phenylalanine. After 3 min of incubation, phenylalanine incorporation occurred only on 100S ribosomes. After 15 min of incubation, insoluble protein, presumed to be polyphenylalanine, and polyphenylalanine associated with both 100S and 70S ribosomes had been formed. These data are in complete accord with the findings of Risebrough, Tissières and Watson (21), who have reported that rapidly synthesized RNA in intact cells becomes associated first with 100S ribosomes.

MECHANISM OF POLYPHENYLALANINE SYNTHESIS

C^{14}-phenylalanine-sRNA was prepared enzymatically, and C^{14}-phenylalanine was found to be transferred to protein without excessive dilution in the presence of a large pool of unlabeled phenylalanine. The transfer required poly U, ribosomes, purified transfer enzyme (16), GTP, and a GTP-generating system (19). The function of GTP is unknown, and undoubtedly several enzymes are required for the over-all synthesis. Such experiments demonstrate that sRNA is an intermediate in polyphenylalanine synthesis; thus, the initial enzymatic reactions are as follows:

$$\text{L-phenylalanine} + \text{ATP} \xrightleftharpoons[\text{phenylalanine-activating enzyme}]{} \text{AMP-phenylalanine} + \text{P-P} \qquad (1)$$

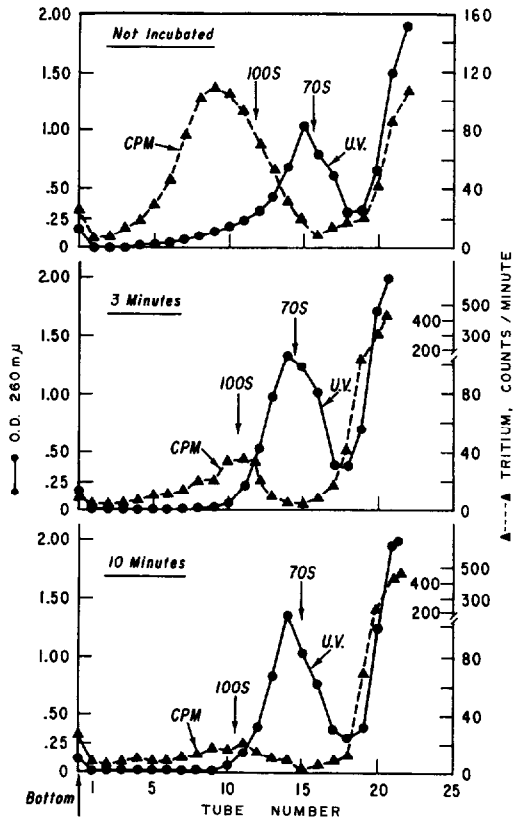


FIG. 7. Association of H^3-poly U with ribosomes. Reactions were begun by addition of 20 μmoles uridylic acid residues in H^3-poly U to each 0.25 ml of reaction mixture. After incubation at 35C for indicated times, samples were briefly chilled in an ice bath and 0.2-ml aliquots were layered on top of sucrose gradients at 3C. Centrifugation and analyses of each fraction were performed as described elsewhere (1).

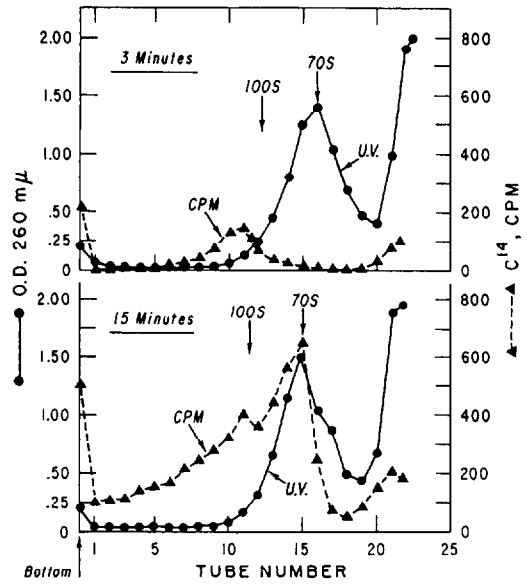


FIG. 8. Incorporation of C^{14}-phenylalanine into protein on ribosomes. A regular reaction mixture was prepared except that C^{14}-phenylalanine (7×10^6 counts/min μmole) was used. 20 mμmoles of uridylic and residue in H^3-poly U were added to 0.25-ml reaction mixtures, which were incubated at 35C for the indicated times. 0.2-ml aliquots then were layered on sucrose gradients at 3C. Centrifugation and analyses of each fraction were performed as described elsewhere (1).

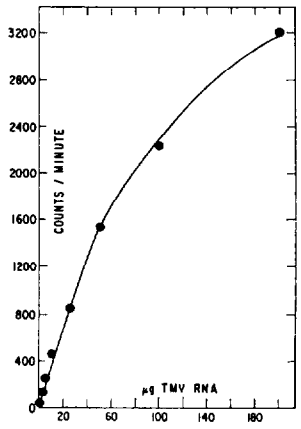


FIG. 9. Stimulation of C^{14}-L-valine incorporation into protein by tobacco mosaic virus RNA. 0.25-ml reaction mixtures were incubated at 37C for 90 min before deproteinization as described elsewhere (19).

$$\text{AMP-phenylalanine} + \text{sRNA} \xrightleftharpoons{\text{phenylalanine-activating enzyme}} \text{phenylalanine-sRNA} + \text{AMP} \quad (2)$$

$$\text{Phenylalanine-sRNA} \xrightarrow{\text{Poly U, ribosomes, transfer enzyme, GTP}} \text{polyphenylalanine} + \text{sRNA} \quad (3)$$

MESSENGER ROLE OF VIRAL RNA

The effects of TMV-RNA in directing protein synthesis in this system were studied in collaboration with Drs. Tsugita and Fraenkel-Conrat of the Virus Laboratory, Berkeley, California (27). The effect of TMV-RNA on C^{14}-valine incorporation into protein is shown in Fig. 9. Proportionality between valine incorporation and TMV-RNA added, in the proper range, was demonstrated. Part of the C^{14}-protein formed a specific precipitate with TMV antisera. Also, part of the C^{14}-protein formed could be purified from reaction mixtures with unlabeled carrier TMV protein by repeated isoelectric precipitations and DEAE column chromatography. Virus reconstitution experiments showed that the C^{14}-protein product strongly and specifically inhibited virus reconstitution. However, the small amount of virus which did reconstitute in the presence of added unlabeled TMV protein and RNA contained small but significant amounts of C^{14}-protein product.

TMV protein is an excellent one to characterize, for its entire amino acid sequence is known and digestion with trypsin converts it into 12 peptides of established amino acid composition and chromatographic properties. The C^{14}-protein synthesized under the direction of TMV-RNA was purified from reaction mixtures with added unlabeled carrier TMV protein and was treated with trypsin. The peptides formed after digestion with trypsin were isolated by Dowex 1 column chromatography. Each peptide was subjected to amino acid analysis and its radioactivity was counted.

The data of Table 6 summarize a small portion of these results (27). Seventeen peptides were isolated and ana-

TABLE 6. *Analysis of peptides obtained from C^{14}-protein synthesized in the presence of either C^{14}-phenylalanine or C^{14}-tyrosine under direction of TMV-RNA*

Peptide No.	Amino Acid Composition	C^{14}-Phenylalanine, counts/min	C^{14}-Tyrosine, counts/min
2	$Thr_{0.9}$ $Glu_{1.0}$ $Val_{1.0}$ (Val)(Arg)	8	0
3	$Thr_{0.1}$ $Ser_{1.4}$ $Glu_{3.0}$ $Pro_{2.1}$ $Val_{2.7}$ *$Phe_{1.0}$* (Try, Lys, Arg)	73	5
11	$Asp_{1.0}$ $Thr_{0.9}$ $Ser_{0.8}$ $Gly_{2.0}$ *$Tyr_{1.0}$* (Arg)	10	44
5	$Val_{1.0}$ *$Tyr_{1.0}$* (Arg)	3	39
8	$Asp_{2.7}$ $Thr_{3.9}$ $Glu_{3.7}$ $Pro_{1.2}$ $Ala_{2.9}$ $Val_{0.6}$ $Ileu_{0.7}$ $Leu_{1.0}$ (Ileu, Arg)	3	11
4	$Asp_{2.0}$ $Ser_{0.9}$ $Pro_{1.0}$ *$Phe_{1.8}$* (Lys)	96	

Reaction mixtures, described in detail elsewhere, contained TMV-RNA and either C^{14}-phenylalanine or C^{14}-tyrosine. After a 90-min incubation at 37 C, ribosomes were removed by centrifugation, and the 100,000 × *g* supernatant solutions were dialyzed extensively. Unlabeled, carrier TMV protein was added and then was purified from supernatant solutions by repeated isoelectric precipitations and DEAE column chromatography. Purified TMV protein was digested with trypsin and 17 peptides were separated by Dowex 1 column chromatography. Amino acid content and radioactivity of each peptide were determined. When yeast messenger RNA was added to reaction mixtures in place of TMV-RNA, the peptides did not contain significant radioactivity. In some cases, peptides were purified further by paper electrophoresis. Data concerning only 6 typical TMV-peptides are presented here. See ref. 27 for complete details.

lyzed; however, space permits presentation of only part of the data. Each peptide in Table 6 is a characteristic TMV peptide. TMV-RNA directed C^{14}-phenylalanine but not C^{14}-tyrosine into phenylalanine-containing peptides; whereas TMV-RNA directed C^{14}-tyrosine, but not C^{14}-phenylalanine, into tyrosine-containing peptides. Although the counts were low, duplicate experiments gave similar results. In addition, several peptides were purified further by paper electrophoresis and still were found to contain radioactivity. The N-terminal amino acid of TMV protein is acetylated and the N-terminal peptide did not contain the expected C^{14}-tyrosine, possibly because acetyl CoA was not added to reaction mixtures. This is the only chemical difference detected so far. Thus, it appears justifiable to conclude that TMV-RNA directs the synthesis of a nonacetylated protein similar to TMV protein in this cell-free *E. coli* system. Such results strongly suggest that a large part of the code may be universal.

SUMMARY

Messenger RNA fractions were found to direct protein synthesis in cell-free *E. coli* extracts. This system afforded a sensitive assay for natural and synthetic messenger RNA and was used both to determine characteristics of messenger RNA and to decipher part of the genetic code.

Polyuridylic acid specifically directed the synthesis of polyphenylalanine via an sRNA intermediate. The activity of poly U varied with its molecular weight; longer polynucleotide chains were more active than shorter ones. Evidence was presented which suggested that each poly U molecule acted catalytically and directed the synthesis of more than one molecule of polyphenylalanine. Single-stranded polynucleotides were shown to have messenger RNA activity but not double- or triple-stranded polymers. Polyphenylalanine synthesis was repressed specifically by poly A, which forms hydrogen-bonded helices with poly U. Longer oligo A chains formed more stable helices with poly U than shorter chains and thus repressed polyphenylalanine synthesis more effectively.

Addition of poly U to reaction mixtures resulted in the formation of polydisperse aggregates of ribosomes sedimenting faster than 70S. Polyphenylalanine was formed initially only on 100S ribosomes. Poly U was degraded rapidly in extracts to mononucleotides, and such degradation was not dependent on protein synthesis.

Randomly ordered polynucleotides were used to direct many amino acids into protein, and the nucleotide compositions of RNA code words corresponding to 15 amino acids were determined. Two coding units corresponding to leucine were found; thus, part of the code was shown to be degenerate.

The number of nucleotides per code word, overlapping and nonoverlapping codes, and the probable extent of code degeneracy were discussed. Tobacco mosaic virus RNA was found to direct the cell-free synthesis of a protein resembling tobacco mosaic virus protein; thus, at least part of the genetic code appears to be universal.

REFERENCES

1. BARONDES, S. H., AND M. W. NIRENBERG. *Science* 138: 810, 1962.
1a. BASILIO, C., A. J. WAHBA, P. LENGYEL, J. F. SPEYER, AND S. OCHOA. *Proc. Natl. Acad. Sci. U. S.* 48: 613, 1962.
2. BRENNER, S. *Proc. Natl. Acad. Sci. U. S.* 43: 687, 1957.
3. CHARGAFF, E., R. LIPSHITZ, C. GREEN, AND M. E. HODES. *J. Biol. Chem.* 192: 223, 1951.
4. FREESE, E. *Brookhaven Symp. Biol.* 12: 63, 1959.
5. GAMOW, G. *Nature* 173: 318, 1954.
6. GIERER, A., AND K. W. MUNDRY. *Nature* 182: 1457, 1958.
7. GRUNBERG-MANAGO, M., AND S. OCHOA. *J. Am. Chem. Soc.* 77: 3165, 1955.
8. KAMEYAMA, T., AND G. D. NOVELLI. *Biochem. Biophys. Res. Commun.* 2: 393, 1960.
9. LAMBORG, M. R., AND P. C. ZAMECNIK. *Biochim. Biophys. Acta* 42: 206, 1960.
10. LENGYEL, P., J. F. SPEYER, C. BASILIO, AND S. OCHOA. *Proc. Natl. Acad. Sci. U. S.* 48: 282, 1962.
11. LENGYEL, P., J. F. SPEYER, AND S. OCHOA. *Proc. Natl. Acad. Sci. U. S.* 47: 1936, 1961.
12. LEVINE, P. A., AND R. S. TIPSON. *J. Biol. Chem.* 111: 313, 1935.
13. MARTIN, R. G., J. H. MATTHAEI, O. W. JONES, AND M. W. NIRENBERG. *Biochem. Biophys. Res. Commun.* 6: 410, 1962.
14. MATTHAEI, J. H., O. W. JONES, R. G. MARTIN, AND M. W. NIRENBERG. *Proc. Natl. Acad. Sci. U. S.* 48: 666, 1962.
15. MATTHAEI, J. H., AND M. W. NIRENBERG. *Proc. Natl. Acad. Sci. U. S.* 47: 1580, 1961.
16. NATHANS, D., AND F. LIPMANN. *Proc. Natl. Acad. Sci. U. S.* 47: 497, 1961.
17. NIRENBERG, M. W., AND J. H. MATTHAEI. *Intern. Congr. Biochem., 5th, Moscow, 1961.*
18. NIRENBERG, M. W., AND J. H. MATTHAEI. *Proc. Natl. Acad. Sci. U. S.* 47: 1588, 1961.
19. NIRENBERG, M. W., J. H. MATTHAEI, AND O. W. JONES. *Proc. Natl. Acad. Sci. U. S.* 48: 104, 1962.
20. NISMAN, B., AND H. FUKUHARA. *Compt. Rend.* 249: 2240, 1959.
21. RISEBROUGH, R. W., A. D. TISSIÈRES, AND J. D. WATSON. *Proc. Natl. Acad. Sci. U. S.* 48: 430, 1962.
22. SPEYER, J. F., P. LENGYEL, C. BASILIO, AND S. OCHOA. *Proc. Natl. Acad. Sci. U. S.* 48: 63, 1962.
23. SPEYER, J. F., P. LENGYEL, C. BASILIO, AND S. OCHOA. *Proc. Natl. Acad. Sci. U. S.* 48: 441, 1962.
24. SPIEGELMAN, S. *Intern. Congr. Biochem., 3rd*, Brussels, 1955, p. 185.
25. TISSIÈRES, A., D. SCHLESSINGER, AND F. GROS. *Proc. Natl. Acad. Sci. U. S.* 46: 1450, 1960.
26. TSUGITA, A., AND H. FRAENKEL-CONRAT. *J. Mol. Biol.* 4: 73, 1962.
27. TSUGITA, A., H. FRAENKEL-CONRAT, M. W. NIRENBERG, AND J. H. MATTHAEI. *Proc. Natl. Acad. Sci. U. S.* 48: 846, 1962.
28. WATSON, J. D., AND F. H. C. CRICK. *Nature* 171: 737, 1953.
29. WILKINS, M. H. F., A. R. STOKES, AND H. R. WILSON. *Nature* 171: 738, 1953.
30. WITTMANN, H. G. *Naturwissenschaften* 48: 729, 1961